



HAL
open science

The evolution of bare nouns in the history of French. The view from calibrated corpora

Pierre Larrivée, Mathieu Goux

► To cite this version:

Pierre Larrivée, Mathieu Goux. The evolution of bare nouns in the history of French. The view from calibrated corpora. *Journal of French Language Studies*, 2024, 28 p. 10.1017/S0959269524000061 . hal-04518628

HAL Id: hal-04518628

<https://hal.science/hal-04518628v1>

Submitted on 15 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ARTICLE

The evolution of bare nouns in the history of French. The view from calibrated corpora

Pierre Larrivée  and Mathieu Goux 

Laboratoire CRISCO, Département des Sciences du Langage, UFR Humanités et Sciences Sociales, Université de Caen Normandie, Normandie Université, Esplanade de la Paix, CS 14032, 14032 Caen cedex 5, France

Corresponding author: Pierre Larrivée; Email: Pierre.Larrivee@Unicaen.fr

(Received 29 May 2023; revised 31 January 2024; accepted 05 February 2024)

Abstract

This article investigates the evolution of bare nouns, used without a determiner, through the history of the French language. The loss of bare nouns is charted through calibrated corpora of non-fictional prose texts from the same genres and region, ranging from the 12th to the 19th century. The change is first completed with nouns in subject function, significantly advances with direct objects, and progresses with obliques. The extensive quantitative documentation demonstrates that the change is impacted by the syntactic function of the noun, along the Accessibility Hierarchy. The speculation is examined that the more accessible functions encourage expression of (definite) determiners, thus explaining the pattern of change.

Keywords: language change; bare nouns; determiners; French; accessibility hierarchy

Résumé

Cet article étudie l'évolution des noms nus, utilisés sans déterminant, à travers l'histoire de la langue française. La perte des noms nus est retracée à travers des corpus calibrés de textes en prose non fictionnelle provenant des mêmes genres et régions, allant du XII^e au XIX^e siècle. Le changement s'achève d'abord avec les noms en fonction sujet, progresse de manière significative avec les objets directs et progresse avec les obliques. La vaste documentation quantitative démontre que le changement est influencé par la fonction syntaxique du nom, le long de la hiérarchie d'accessibilité. L'hypothèse est examinée selon laquelle les fonctions les plus accessibles encouragent l'expression de déterminants (définis), expliquant ainsi le modèle de changement.

Mots-clés: changement linguistique; noms nus; déterminants; français; hiérarchie d'accessibilité

1. INTRODUCTION

One notable grammatical change in the history of French is the loss of bare nouns. While it was still possible to use bare nouns in Medieval French, this possibility seems to have been reduced over time, especially as compared with other Romance languages (Carlier and Lamiroy 2018a, 2018b; for Germanic, see Skrzypek, Piotrowska and Jaworski 2021). Although quantitative information is provided in some studies (notably Simonenko and Carlier 2020b for the period before 1400, and Simonenko and Carlier 2020a before 1700), the general progression of the loss remains to be charted, in terms of e.g. syntactic and semantic correlative factors.

Demonstrating the impact of syntactic function on bare noun loss is what this article aims to do. It uses corpora of non-fictional prose calibrated for genres (three text types from the legal and historiographical domains) and region (Normandy) together extending over 8 centuries (from the 12th to the 19th century) to do so. Function is shown to correlate to the bare noun loss, from central to peripheral functions, alongside the Accessibility Hierarchy (Keenan and Comrie 1977).

First, the article discusses studies of bare noun evolution. The method by which a systematic investigation of the evolution is conducted is then explained. The rates of use of bare nouns in the selected diachronic corpora are then presented; and the syntactic pattern of change is discussed and speculated upon. The proposals are summarized in the conclusion, which identifies issues for future research.

2. STATE OF THE ART

The NP¹ generally contains a determiner in contemporary French (Abeillé and Godard 2021: chapter V). There are no argumental² contexts of use (as is the case with English bare plurals, or Italian bare mass nouns) in which a French NP can commonly be found without a determiner (Lauwers 2012). This contrasts with other Romance languages: Carlier and Lamiroy (2018a), who study the emergence of the determiner category (see also Combettes 2001), cite figures of 20 % of bare nouns in Spanish and 15% in Italian, as compared to 6% in contemporary French, where bare nouns are restricted to specific cases such as proper nouns (1) and nouns with metalinguistic reference (2).

- (1) **Paul** est plombier.
Paul be-PR-3S plumber
“Paul is a plumber. »
- (2) “**Paul**” est un nom propre. ”
Paul be-PR-3S a noun proper
““Paul” is a proper noun.”

Specific constructions of higher registers, such as coordination illustrated by (3) (Roodenburg 2004, Märzhäuser 2013, Riegel *et al.* 2014: 308–315), and idiomatic

¹The term NP is here used in a spirit of theoretical neutrality.

²Non-argumental environments do host bare nouns in French, mainly appositions (6) and predicative complements (7) (de Swart *et al.* 2007, de Swart and Zwarts 2009, Lauwers 2011).

expressions in object (see (4), Chaurand 1991, Gross et Valli 1991) and oblique functions (see (5), Anscombe 1991a and 1991b, Vigier 2017) also allow nouns to be used without a determiner.

- (3) ^G **Femmes et enfants** sont les premières victimes
 Women and children be-PR-3P the first victims
 “Women and children are the first victims.”
- (4) Elle a perdu **pied**.
 3S-NOM have-PR-3S lost-PTCP foot
 “She lost balance.”
- (5) Il est **en colère**.
 3S-NOM is-PR-3S in anger
 “He’s angry.”

And so do bare nouns in appositive and predicative functions (Beyssade 2011, Lauwers 2014).

- (6) ^G Maria, **femme de ménage**, est au chômage
 Maria, woman of cleaning, be-PR-3S at-the unemployment
 “Maria, a cleaning-lady, is unemployed.”
- (7) ^G Il est **éboueur** depuis douze ans.
 3S-NOM be-PR-3S garbage-collector since twelve years
 “He’s been a garbage-collector for twelve years.”

The suggestion is that across these constructions, bare noun relates to a property rather than a delimited entity (notably Anscombe 1991a, 1991b): it is not a particular foot that is evoked in (4), but the notion of one’s footing. Property denotation is a recurring factor of bare noun use. However, other factors must also be involved, since proper nouns, which are generally used without a determiner, definitely have a referential value.

How did bare nouns in contemporary French get restricted to a disparate set of environments? The rates, functions, and semantic values discussed by diachronic studies are evoked in the following paragraphs.

It is useful to remember that bare nouns are thought to be the default case in Latin (Pinkster 2015: 48–50): Carlier and Lamiroy (2018a) cite a figure of 77%. Although the language has possessive, demonstrative and numeral items with the noun, they are considered akin to modifiers by Sornicola (2009: 33–34) and are thought to gradually evolve as members of the determiner class in Gallo-Romance.

Rates of bare nouns over the history of French are provided by various studies (see Sakari 1988, Haussalo 2014). Carlier and Lamiroy (2018a) propose 32% in Old French and, as mentioned, 6% in Modern French. Reporting on two translations of Cicero’s *De Inventione* in 1282 by Jean d’Antioche and in 1932 by de Bornecque, and where the Latin original contains 86,66 % of bare nouns, Goyens (1994) identifies figures of 40,76 % in Old French and 15,98 % in contemporary French.

Similarly, Saikali-Sleiman (2006)'s dissertation studies the distribution of bare nouns in three successive versions of the *Conquête de Constantinople* (by Villehardouin at the beginning of the 13th century, in the 1585 translation by Blaise de Vigenère and in the 1939 translation by Edmond Faral). The overall proportions are at 33%, 40% and 27%, with an unexpected rise in the 16th century,³ and surprisingly little change between the initial and the final period. More extensive quantitative documentation is provided by Simonenko and Carlier (2020b: 221). Banking on the MCVF, they show that bare nouns are receding from the 1100 to the 1600. However, the progression is irregular, with an unexpected higher rate of bare nouns in the 1500 than in the 1400 in subject position, as well as in the 1200 and the 1500 as compared to the 1100 and the 1400 in object position (see their Figure 11). This is likely due to the heterogeneous nature of the corpus (on regularity provided by homogeneous corpora, see the discussion and results in Vigier 2015, 2017).

Regarding function, in Old French, bare nouns are found in all of them (Marchello-Nizia 1999: 76), although in no function are they the default option (Marchello-Nizia *et al.* 2020: 973). Saikali-Sleiman (2006) notes that real decline⁴ is manifested with subjects (53%, 29% and 18%) and direct objects (38%, 33%, 29%). No significant change is however noted after prepositions (mostly concerning à “to”, de “from”, en “in” and pour “for”) that go from 29% to 43% and 28%.⁵ Some further quantitative data is provided by Marchello-Nizia *et al.* (2020: 972–973), who cite 3% of bare nouns in subject position at the end of the 15th century (in a historical text by Commynes), and 0% by the middle of the 16th century (in Calvin's sermons).⁶ Simonenko and Carlier (2020a) contrast bare nouns in subject position (going from about 25%, to 20%, to 18%, to 15%, to 18% and 15% for the 12th, 13th, 14th, 15th, 16th and 17th, judging from their Figure 11) and in object position (from 40% downward to about 35%). The progression seems however more conservative than has been observed in other studies, calling for complementary studies. A similar sensitivity to syntactic function has also been observed for partitive determiners by Carlier and Lamiroy (2018b). Non-quantitative approaches note that bare nouns in object position are increasingly found in set expressions⁷ through time (Gross and Valli 1991).

The semantic contribution of bare nouns relates to lack of individuation of the NP's referent according to Foulet:

³Which could be attributed to the archaic natures of the literary genre of the military chronicle; especially for the 16th French century, known for its elevated style in imitation of antiquity models (see Lorian 1973).

⁴Also noted are changes with nouns preceded by adjective *grand* “big” (41%, 31%, 28%), and attributive and appositive complements (with aggregated figures of 36%, 35%, 29%).

⁵See also Vigier (2017) for detailed data regarding bare nouns with prepositions *en*, *dans* and *dedans* from the 16th century.

⁶They also note 11% of bare nouns in direct object function in the middle of the 20th century (Sartre's *Letters*); and 58% of indirect object bare nouns for archaic French 9th century *Cantilène de Sainte-Eulalie*.

⁷Since, as pointed out by a reviewer points, set phrases can comprise structures from older states of languages.

si l'individualité ne ressort pas nettement, si nous avons affaire à un type plutôt qu'à un individu, ou si l'individu nous est présenté comme devant satisfaire à telles ou telles conditions qui pourront être ou ne pas être remplies⁸ (Foulet 1929: 56)

This seems convergent with the cases listed by Buridant (2019: 147–153). However, disagreement as to the property-denoting reference of bare nouns is expressed by Marchello-Nizia *et al.* (2020: 973) for whom there is no unique meaning to the absence of determiners, since such absence is compatible with proper nouns and nouns evoking a unique referent (8), as well as with generics (9), plural indefinites (10), mass (11) or abstract (12) nouns, following the examples they cite.

- (8) Icele tere, [...] dun il esteit, **soleil** n'i luist
This land of which 3S-NOM be-PA-3S sun NEG there shine-PR-3S
“In the land that he was from, the sun didn't shine.”
 (“au pays dont il vient, le soleil ne brille pas”, *Chanson de Roland*, v. 979–980)
- (9) **In aiudha, per dreit**
“in help, for rights”
 (“par mon aide, par l'équité”, *Serments de Strasbourg*)
- (10) **Colpes non auret**
sin NEG have-PA-3S
“She was without sin.”
 (“elle était sans péché”, *Cantilène de Sainte-Eulalie*)
- (11) Devant eals unt **dulz e blanc pain**
before 3P have-PR-3P soft and white bread
“They have soft and white bread before them.”
 (“devant eux, ils ont du pain doux et blanc”, *Benedeit-Brendan*, v. 699)
- (12) demeinent **grant dolor**
show-PR-3P great pain
“They showed great pain.”
 (“ils manifestent une grande souffrance”, *Chanson de Roland*, v. 2695)

Carlier and Lamiroy (2018a) suggest that the bare nouns are used with singular and plural count nouns in the 14th century, with plural count and mass nouns in the 15th and 16th, before disappearing in the 17th. According to Foulet (1929), the decline is related to bare nouns being challenged by the emergence of the indefinite article, that would have a similar meaning contribution, and would increasingly be found with nouns with an indefinite and abstract interpretation (see also Déchaine *et al.* 2018, and Simonenko and Carlier 2020b).

⁸“If the individuality does not come out strongly, if we are dealing with a type rather than an individual, or if the individual is presented to us as having to satisfy some conditions that may or may not be fulfilled.” See also Riegel *et al.* (2014: 312).

The existing studies support the view that bare nouns are a declining grammatical option, whose productivity is gradually eroded by function, types of nouns, and collocations. Regarding function, it looks as though bare nouns in Medieval French disappear from subject function, decline as direct objects, and regress in obliques, although there are considerable disparities in numbers from homogeneous corpora (such as translations), and variability in heterogeneous databases. As for types of nouns, there is little agreement in the literature, probably due to the difficulty of diagnosing property denotation in actual usage through history. Subsisting uses of bare nouns seem to be increasingly engaging in collocations, although again the identification of collocative uses diachronically raises methodological issues. The first question is what this study of the evolution of bare nouns in Medieval French focuses on. How it intends to proceed is presented in the next section.

3. METHOD AND CORPUS

The purpose of this article is to conduct a systematic investigation of one potential syntactic correlate of the declining use of bare nouns in the history of the French language.

Development is ideally perceptible in data calibrated through the documented period. Such a calibrated corpus is difficult to assemble for several well-known reasons. As we go back in time, the set of available texts is narrower, they tend to be heterogenous in nature, size and geographical origins; texts of a literary genre can be archaic and their stylistic pursuit can make them unreliable witnesses of effective language evolution (Prévost, 2015; Goux and Larrivé, 2020). Fortunately, these difficulties are gradually being resolved by new resources, with tools that support data identification and extraction.⁹

Among these resources, the *MICLE* and *Chroniques* corpora provide calibrated data through time.¹⁰ The objective of their creation was to reduce data heterogeneity. By proposing prose texts written only in Normandy, they contribute to eliminating undue diatopic influences; by each selecting non-fictional texts relating to the same domain of activity, they present data that are mostly exempt from the stylistic pursuit characteristic of literary texts. Such calibrated corpora of non-literary material have been shown to provide a more stable and regular picture of language change (Goux and Larrivé 2020; Goux 2022b).

The texts from *MICLE* belong to two different genres: trial material is listed in Table 1¹¹ and the procedure styles, that explain procedure to be followed in court, are in Table 2.

⁹For a survey of NLP tools and annotation of French corpora, see Gabay *et al.* (2020a,b), and for Old French, see Grobol *et al.* (2022).

¹⁰We gratefully acknowledge the generous funding of the ANR and DFG agencies for the *MICLE* corpus (Projet ANR-20-FRAL-0001), and of the Normandy Region for the *Chronicle* corpus through the HIGH-TECH project (<<http://crisco.unicaen.fr/recherche/projet-rin-high-tech-1089578.kjsp?RH=1531402918899>>).

¹¹All bibliographical information is available on the Crisco-TXM portal: <https://txm-crisco.huma-num.fr/txm/>

Table 1. Trial material from the *MICLE* corpus

Date	Title	Author	Source of the document	Token number
1207	<i>Assises de Normandie</i>	Anon.	<i>Établissements et coutumes</i> , Ed. by A. I. Marnier (1839); available on <i>Gallica</i> ¹²	4 766
1292	<i>Anglo-Norman Year Books (1)</i>	Anon.	<i>Year Books of the Reign of King Edward I</i> , Ed. by A. J. Harwood (1863-1879); available at the British Library ¹³	33 557
1340	<i>Anglo-Norman Year Books (5)</i>	Anon.	<i>Year Books of the Reign of King Edward I</i> , Ed. by A. J. Harwood (1863-1879); available at the British Library ¹⁴	49 203
1431	<i>Procès de Jeanne d'Arc</i>	Anon.	<i>Minute Française du procès de condamnation de Jeanne d'Arc</i> , Ed. by Abbé Dubois (1849)	41 967
1450	<i>Réhabilitation de Jeanne d'Arc</i>	Anon.	<i>Procès de condamnation de Jeanne d'Arc</i> , Ed. by J. Quicherat (1844)	8 716
1502	<i>L'Héritage de Marie de la Roche Guyon</i>	Anon.	Ed. by V. Deplaigne (2009)	41 149
1563	<i>Grefte de procès de sorcellerie à Guernesey</i>	Anon.	List and Index Society from Guernesey transcr. (1696)	12 190
1591	<i>Procès verbal pour délivrer une fille possédée à Louviers</i>	Anon.	Ed. by B. de Moray (1883); available on <i>Gallica</i> ¹⁵	30 773
1643	<i>Procès-verbal de Magdeleine Bavent</i>	Mr le Penitentier d'Evreux	Available on <i>Gallica</i> ¹⁶	3 343
Total				225 664

While the first set runs over five centuries, the second is limited to three, but is concentrated on the period where change is believed to be decisive.

A third genre was added from the corpus *Chroniques*, and the text that were ready to exploit at the time of writing are listed in Table 3.

These cover the very beginning, middle and end of the documented period. Although offering three different types of temporal distribution, using three text types has the added benefit of controlling for outlier results and distortions that may be due to a specific text, or text types.

Once the texts had been digitized and transcribed, tokens were lemmatized and tagged following the Universal Dependencies guidelines¹⁷ with the HOPS syntactic parser (Grobol and Crabbé 2021). The parser first identifies the POS (“part of

¹²<https://gallica.bnf.fr/ark:/12148/bpt6k57063292>

¹³<http://explore.bl.uk/BLVU1:LSCOP-ALL:BL01001077075>

¹⁴<http://explore.bl.uk/BLVU1:LSCOP-ALL:BL01001077075>

¹⁵<https://gallica.bnf.fr/ark:/12148/bpt6k77217c>

¹⁶<https://gallica.bnf.fr/ark:/12148/bpt6k84857r>

¹⁷<https://universaldependencies.org/>

Table 2. Procedure styles from the *MICLE* corpus

Date	Title	Author	Source of the document	Token number
1314	<i>Atiremenz et jugiés d'eschequiers</i>	Anon.	Ed. by R. Génestal and E.-J. Tardif (1924)	15 232
1425	<i>Style et usage de l'Echiquier</i>	Anon.	Ed. by Société des Antiquaires de Normandie (1851), available on <i>Gallica</i> ¹⁸	71 903
1578	<i>Commentaire du droit civil tant public que privé observé en Normandie</i>	Terrien	pp. 339 to 402 (ninth book); available on <i>Gallica</i> ¹⁹	42 800
Total				129 935

Table 3. Material from the *Chroniques* corpus

Date	Title	Author	Origin of the document:	Token number
1190	<i>History of the Kings of England</i>	Anon.	“An Early French Prose History of the Kings of England”, Ed. by D. B. Tyson (1975)	5 801
1487	<i>Crocniques de Normandie</i>	Anon.	Ed. by A. Hellot (1881)	45 958
1862	<i>Histoire de la Normandie ancienne et moderne</i>	Ed. by C.Barthélémy	Available on <i>Gallica</i> ²⁰	50 539
Total				102 298

speech”) of the token; then the dependency relations between tokens and the nature of that dependency relation, such as subject, direct object, oblique.²¹ The files generated by the HOPS parser (known as a CONLL-U files) were then converted into XML-TEI files through a Python scripts specifically developed for the MICLE project, and uploaded unto the TXM software (Heiden *et al.* 2010), to allow CQL requests to be made.²²

We used the following query to spot the occurrences of bare nouns in the corpora:

```
a:[udpos!="DET"][udpos!="DET"]{0,2}[udpos="NOUN" & function="xxx"
& n=a.head]
```

¹⁸<https://gallica.bnf.fr/ark:/12148/bpt6k200065b>

¹⁹<https://gallica.bnf.fr/ark:/12148/bpt6k9107304x>

²⁰<https://gallica.bnf.fr/ark:/12148/bpt6k5781079k>

²¹For a complete list of the dependency relations used in the UD systems, see <https://universaldependencies.org/u/dep/index.html>.

²²<https://inl.github.io/BlackLab/guide/corpus-query-language.html>

The attribute “udpos” allow us to search for a specific POS in the UD tagset;²³ and the “!=” to exclude the POS in question, and return all other POS. The accolades {0,2} define an interval between the head of the NP and the noun, and the function of the NP is captured by the tag “function”, whether it be nominal subject (nsubj), direct object (obj) and oblique (obl). The initial element of the request, “a:” captures all the information relating to the first member of the NP, and ensures that the search is done within the boundaries of the NP (as defined by the dependents of the head noun).

Once the occurrences have been automatically retrieved using TXM, we examined them manually to eliminate false positives (mainly due to residual tagging errors in the corpora). The general principle for bare noun identification was that a determiner such as the (in)definite article could have been used at that time but was not. That is why proper nouns were left out of the searches, as they are but rarely used with a determiner. For the same reason, quantified phrases such as *beaucoup de* “a lot of N” that came out in the searches were not counted as bare noun. However, since a determiner can precede the noun in those structures, a bare noun after a partitive in a negative context (*pas de N*, “no N”) was.

Having explained the methodological choices of corpus selection and data extraction, we now turn to the evolution in the rates of use of bare nouns with regards to the central factor of syntactic functions.

4. BARE NOUNS LOSS IS IMPACTED BY SYNTACTIC FUNCTION

4.1 Bare nouns in subject function

We start by presenting the results of the investigation of the evolution of Medieval French bare nouns through time in the subject function, where it is believed to be first lost.

Rates of subject bare nouns in the trial material of the *Micle* corpus, as compared to the sum of nouns without and with determiners, are presented in Table 4.

The identified proportions represent a decline. They go from over 4% at the beginning of the 13th century, to around 1% from the end of the 13th to the beginning of the 16th century, after which subject bare nouns go out of use. Example (13) provides one of the earliest instance of bare nouns in the corpus ; (14), one of the latest.

(13) E se **pès** est fete, ... (1207, *Assises*)
 And if **peace** be-PR-3S do-PTCP,
 “And if peace is made, ...”

(14) et denssoit comme font **jeunes femmes** (1502, *Roche-Guyon*)
 and dance-PA-3P like do-PR-3P **young women**
 “and danced like young women do”

The intermediate period is further documented by the procedure style texts, see Table 5.

²³<https://universaldependencies.org/u/pos/index.html>

Table 4. Numbers and rates of subject bare nouns in *MICLE* trial material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1207 Assises Normandie</i>	2	47	4,1%
<i>1292 ANYB</i>	5	556	0,9%
<i>1340 ANYB</i>	12	1090	1,1%
<i>1431 Procès Jeanne d'Arc</i>	3	433	0,7%
<i>1450 Rehabilitation Jeanne d'Arc</i>	0	87	0%
<i>1502 Roche-Guyon</i>	5	714	0,7%
<i>1563 Guernesey</i>	0	242	0%
<i>1591 Fille possedee</i>	0	278	0%
<i>1643 Procès Bavent</i>	0	36	0%
Totals	27	3483	0,8%

Table 5. Numbers and rates of subject bare nouns in *MICLE* procedure material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1314 Attiremens et Jugies</i>	3	473	0,6%
<i>1425 Style et Usage Echiquier</i>	11	1762	0,6%
<i>1578 Terrien²⁴</i>	4	1013	0,4%
Totals	18	3248	0,6%

The proportions below 1% are in keeping with the trial figures, that are around 1%. Once again, we illustrate with one of the earliest (15) and latest (16) instances of bare nouns in the procedure material.

- (15) *quar simple plevine ne descent pas jusques a l'eir* (1314, *Attiremens*)
 because **simple pledge** NEG descend-PR-3S NEG until at the heir
 “because simple pledge doesn’t descend up to the heir”

²⁴In total, we retrieved 28 bare nouns in subject function for this text, but 24 of these are of the metalinguistic type, as per example (2) above and as illustrated below. These have been excluded, as were proper nouns from all texts.

COUR est une assemblée de sages hommes
 COURT be-PR-3S an assembly of wise men
 “A COURT is an assembly of wise men” (*Terrien*, p. 339)

- (16) par lesquels **droit** doit estre fait... (1578, *Terrien*)
 by which **law** must-PR-3S be-INF do-PTCP
 “for which the law must be done...”

The overall evolution is confirmed by the figures of the *Chroniques* corpus (Table 6).

Table 6. Numbers and rates of subject bare nouns in *Chronique* material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1190 Kings of England</i>	6	89	6,3%
<i>1487 Cronicque</i>	12	946	1,3%
<i>1862 Histoire de la Normandie ancienne</i>	1	1039	0,1%
Totals	19	2074	0,9%

- (17) si **grant anguisse** fu après (1190, *Kings*)
 if **great anguish** be-PA-3S after
 “if a great anguish came after...”

- (18)²⁵ En 1119, **diverses affaires** eurent lieu... (1862, *Histoire*)
 In 1119, **several issues** have-PA-3P location...
 “In 1119, several issues arose...”

The shape of the evolution of bare nouns in the three corpora is presented in the following graph (Fig. 1). The texts are tied to the nearest half-century (and to the 1400 point for 1425 procedure style; and averaging for the 1450 point the results of the 1431 and 1450 trials), the rates in percentages are rounded to the nearest point (except for the 1450 and the 1850 point, which would not have been represented had they not be rounded at 0.5 and 0.1% respectively).

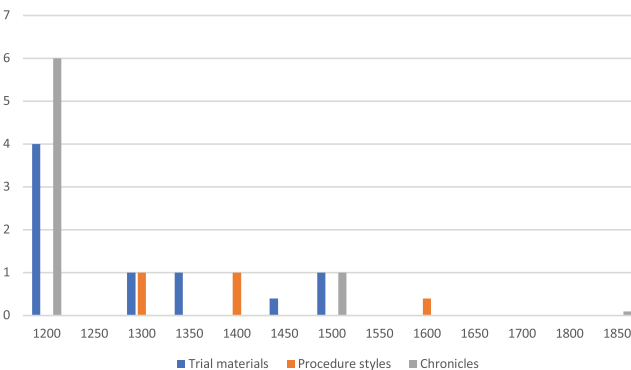


Figure 1. Evolution of subject bare nouns in three sub-corpora.

²⁵This is the only occurrence found in the later chronicle. On the one hand, as an article could be added to the subject but was not, it formally counts as a bare NP. On the other, it is a matter of appreciation whether indefinite “diverses” could not be considered as a kind of determiner. We leave the question open.

It shows a use at around 5% in the earliest period, and a subsequent rate around 1% until the use disappears at the end of the 16th century.

4.2 Bare nouns in object function

The same quantitative analysis was performed for bare nouns in object position.

The numbers in the *MICLE* trial material corpus are presented below (Table 7).

Table 7. Numbers and rates of object bare nouns in *MICLE* trial material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1207 Assises Normandie</i>	0	37	0 %
<i>1292 ANYB</i>	47	593	7,3 %
<i>1340 ANYB</i>	46	1011	4,4 %
<i>1431 Proces Jeanne d'Arc</i>	82	634	11,5 %
<i>1450 Rehabilitation Jeanne d'Arc</i>	17	133	11,3 %
<i>1502 Roche-Guyon</i>	47	588	7,4 %
<i>1563 Guernesey</i>	13	226	5,4 %
<i>1591 Fille possedee</i>	28	650	4,1 %
<i>1643 Procès Bavent</i>	3	98	3 %
Totals	283	3970	6,7 %

(19) si yl aveyt **bone seisine** ou nun (1292, ANYB)
 if 3S-NOM have-PA-3S **good possession** or not
 “if he had rightful possession or not”

(20) elle ne fit pas **grande resistance** (1643, *Proces Bavent*)
 3S.Fem-NOM NEG do-PA-3S NEG **great resistance**
 “she didn’t resist much”

The trial material does not present the expected overall sharp decline, but rather a rise to around 11% in the 1430–1450 period, followed by a slow decrease. The impression of a rise is partly caused by the absence of bare noun objects in the first text, and a lower rate in the 1340 text than either the preceding or following text; as well as by the relatively high rates in the two 15th-century texts. Whether this latter observation is due to those particular texts, or to the period, can be established by looking at the same phenomenon in the procedure style corpus. The figures are as follows (Table 8):

Table 8. Numbers and rates of object bare nouns in the *MICLE* procedure material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
1314 <i>Attiremens et Jugies</i>	31	387	7,4%
1425 <i>Style et Usage Echiquier</i>	72	1447	4,7%
1578 <i>Terrien</i>	64	1084	5,6%
Totals	167	2918	5,4%

- (21) et nul en Normandie n'a **plaine garde** (1314, *Atiremens*)
and no one in Normandy NEG have-PR-3S **full custody**
“and no one in Normandy has full custody”
- (22) [la chose] requiert **prompte prouision** (1578, *Terrien*)
[the matter] require-PR-3S **quick provision**
“the matter requires a quick provision”

They are more in keeping with the figures in the trial corpus, suggesting that the observed peak may be an effect of the 1431 and 1450 trials. Yet a more nuanced picture is provided by the *Chroniques* corpus, whose numbers are given in the next table (Table 9):

Table 9. Numbers and rates of object bare nouns in *Chroniques* material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
1190 <i>Kings of England</i>	12	119	9,2%
1487 <i>Cronicque</i>	68	754	8,3%
1862 <i>Histoire de la Normandie ancienne</i>	37	1472	2,5%
Totals	117	2345	4,8%

- (23) si menat **seinte vie** pus (1190, *Kings*)
then lead-PA-3S **holy life** then
“he then led a holy life”
- (24) mais il fit toujours si **bonne contenance** (1862, *Histoire*)
but 3S-NOM do-PA-3S always so **good composure**
“but he always had such a good composure”

The 15th century chronicle displays a higher rate than the earlier 1425 style, but lower than the two 15th century trials. This is shown by the following graph (Fig. 2), which provides a representation of object bare noun rates in the three corpora.

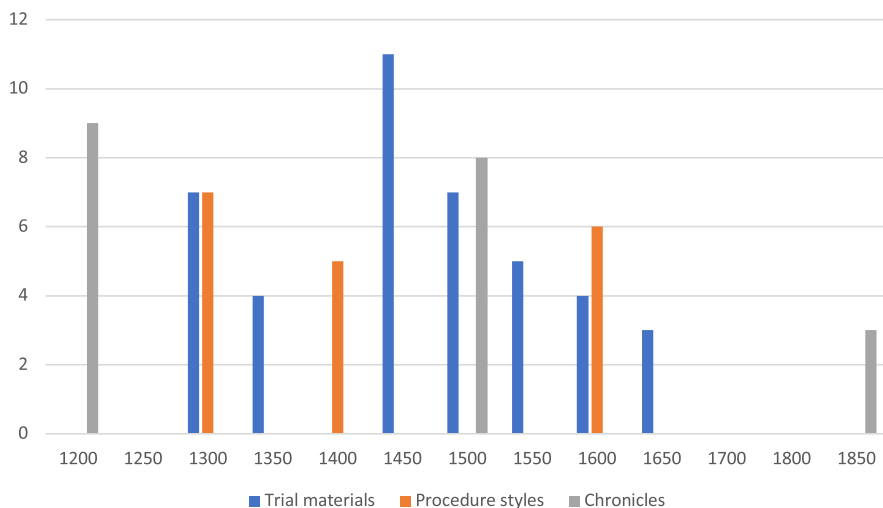


Figure 2. Evolution of object bare nouns in three sub-corpora.

Overall, the loss is one that goes from nearly 10% to 3%, with some variance in the period between 1430 and 1550. The development of bare nouns in the last investigated syntactic category is considered in the next section.

4.3 Bare nouns in oblique function

The annotation tool used separates objects from obliques. Obliques comprise a “nominal (noun, pronoun, noun phrase) functioning as a non-core (oblique) argument or adjunct. This means that it functionally corresponds to an adverbial attaching to a verb, adjective or other adverb”, according to the UD guidelines.²⁶ Note that except in the case of clitics, indirect objects are also annotated as obliques, an issue on which we come back later in this section.

The nouns in oblique function without determiners and with determiners have been extracted and manually examined in each of the trial texts of the *MICLE* corpus for reliable figures to be established. The proportion of nouns without determiners to the sum of nouns with and without was then calculated. The results are presented in the table below (Table 10):

²⁶<https://universaldependencies.org/u/dep/obl.html>

Table 10. Numbers and rates of oblique bare nouns in *MICLE* trial material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
1207 <i>Assises Normandie</i>	35	57	38%
1292 <i>ANYB</i>	597	1270	32%
1340 <i>ANYB</i>	1086	1809	37,5%
1431 <i>Proces Jeanne d'Arc</i>	378	1448	21%
1450 <i>Rehabilitation Jeanne d'Arc</i>	106	356	22,9%
1502 <i>Roche-Guyon</i>	360	1794	16,7%
1563 <i>Guernesey</i>	227	413	35,5%
1591 <i>Fille possedee</i>	304	1250	19,6%
1643 <i>Proces Bavent</i>	41	167	19,7%
Totals	3134	8564	26,8%

- (25) Se cil qui ont esté **en jugement** (1207, *Assises*)
 If those who have-PR-3P be-PTCP **in trial**
 “If those who had been on trial...”
- (26) Dagon estoit **en forme** de beste hideuse (1643, *Proces*)
 Dagon be-PA-3S **in shape** of hideous beast
 “Dagon had the shape of a hideous beast”

The change goes from 40% in the period to 1340, to 20% in the texts from 1430 (despite the unexpectedly high proportion in the Guernesey witch trials).

The pattern is more conservative in the procedure texts (Table 11):

Table 11. Numbers and rates of oblique bare nouns in *MICLE* procedure material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
1314 <i>Attiremens et Jugies</i>	266	517	34%
1425 <i>Style et Usage Echiquier</i>	1540	2636	36,9%
1578 <i>Terrien</i>	725	1989	27%
Totals	2531	5142	49,2%

- (27) il le lessa metre **en defaute** (1314, *Attiremens*)
 3S-NOM 3S-ACC let-PA-3S put-INF **in default**
 “he let it go to default”

- (28) celuy qui est trouvé **en jugeme[n]t** (1578, *Terrien*)
 this who be-PR-3S find-PTCP **in trial**
 “the one who is found in trial”

The situation in the chronicles is more in keeping with this trial corpora, although the initial period has slightly lower figures than expected and the later one slightly higher ones (Table 12):

Table 12. Numbers and rates of oblique bare nouns in *Chroniques* material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1190 Kings of England</i>	81	152	34,8%
<i>1487 Cronique</i>	619	2018	23,5%
<i>1862 Histoire de la Normandie ancienne</i>	498	2228	18,3%
Totals	1198	4398	21,4%

- (29) ducs Williame vint e le conquest **a force e armes** (1190, *Kings*)
 duke William come-PA-3S and 3S-ACC conquer-PA-3S **at force and weapons**
 “duke William came and conquered it by force and with weapons”
- (30) Henri gouverna ensuite **en paix** (1862, *Histoire*)
 Henri rule-PA-3S then **in peace**
 “Henri ruled in peace from then on.”

This is shown by the following graph (Fig. 3):

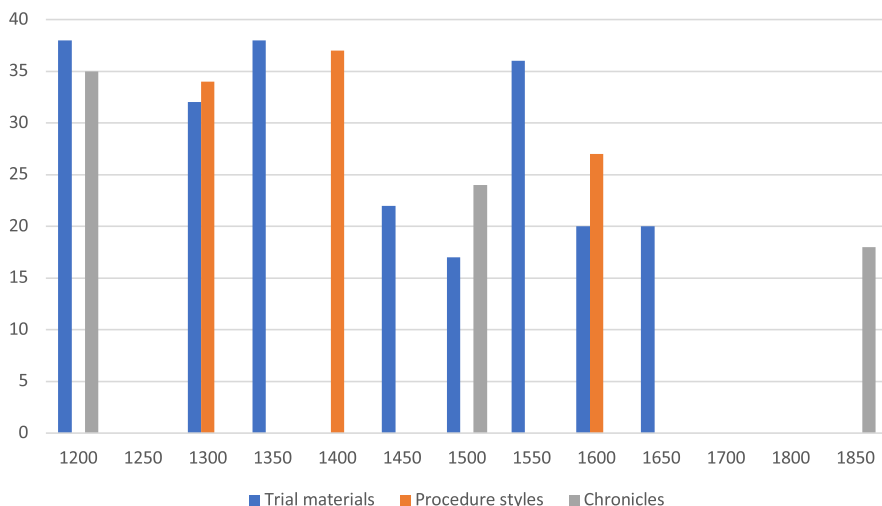


Figure 3. Evolution of oblique bare nouns in three sub-corpora.

The overall course of bare nouns in oblique function is one by which, despite the spike in the Guernsey text and the high proportions provided by procedure texts, the rates go from high-to-mid 30% until 1430, and fall to about 20%, a rate still found in the 19th century according to the last *Chroniques* text.

Despite the shortcoming of the annotations tools, we felt it would be good to be able to distinguish between indirect objects and adjuncts. A strategy was devised to achieve this. We first identified verbs used in conjunction with a clitic pronoun like *lui*, *y*, or *en*, which are correctly recognized as indirect objects (“*ioj*” in the UD tagset) by the HOPS parser. An oblique NP with these specific verbs can thus be safely assumed to be an indirect object, and especially when introduced with prepositions like *à*, *de* or *en*. We then proceeded to count the indirect objects with and without a determiner for each function, with results reported below (Tables 13–15, Fig. 4) :

Table 13. Numbers and rates of indirect objects bare nouns in *MICLE* trial material

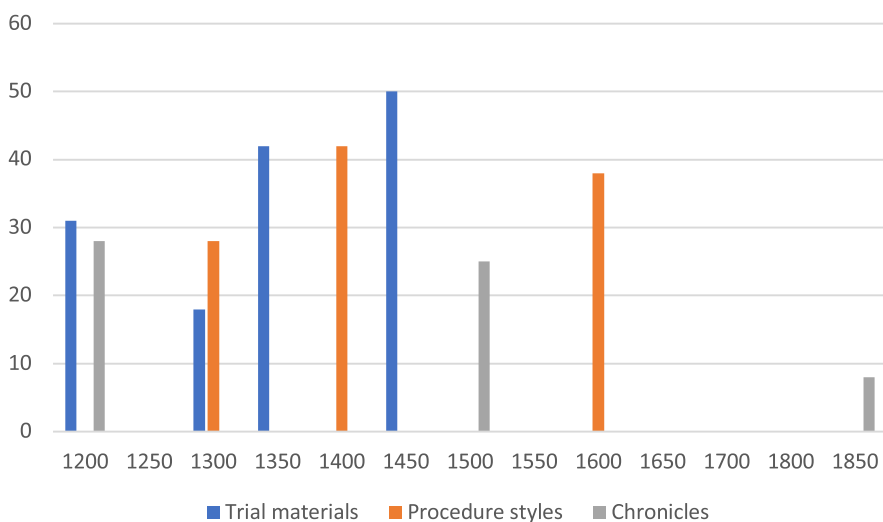
	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1207 Assises Normandie</i>	14	31	31,1%
<i>1292 ANYB</i>	92	148	38,3%
<i>1340 ANYB</i>	135	186	42%
<i>1431 Proces Jeanne d'Arc</i>	74	241	23,5%
<i>1450 Rehabilitation Jeanne d'Arc</i>	9	25	26,5%
<i>1502 Roche-Guyon</i>	0	312	–%
<i>1563 Guernesey</i>	0	65	–%
<i>1591 Fille possedee</i>	0	73	–%
<i>1643 Proces Bavent</i>	0	11	–%
Totals	324	1092	22,9%

Table 14. Numbers and rates of indirect objects bare nouns in *MICLE* procedure material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1314 Attiremens et Jugies</i>	33	85	28%
<i>1425 Style et Usage Echiquier</i>	305	429	41,6%
<i>1578 Terrien</i>	141	234	37,6%
Totals	479	748	39%

Table 15. Numbers and rates of indirect objects bare nouns in *Chroniques* material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1190 Kings of England</i>	7	18	28%
<i>1487 Cronique</i>	123	374	24,7%
<i>1862 Histoire de la Normandie ancienne</i>	19	236	7,5%
Totals	149	628	19,2%

**Figure 4.** Evolution of indirect object bare nouns in three sub-corpora.

- (31) Nus [...] portames un autre bref **a vicomte** (1292 *ANYB*)
 1P [...] bring-PA-1P an other brief at viscount
 “We brought another brief to the viscount”
- (32) lequel fut semblablement délaissé **à justice séculière**
 (1450 *Réhabilitation*)
 who be-PA-3S similarly abandon-PTCP at justice secular
 “who was similarly left to secular justice”
- (33) et il est apporté **a justice** (1314 *Attiremens*)
 and 3S-NOM be-PR-3S bring-PTCP at justice
 “he is brought before justice”

- (34) il doit aller à **Cour** (1578 *Terrien*)
 3S-NOM mus-PR-3St go-INF at court
 “he must go before the court”
- (35) le alme alat a **debles** (1190 *Kings*)
 the soul go-PA-3S at devil
 “the soul went to the devil”
- (36) les Parisiens ne voulant pas
 the Parisians NEG wish-PTCP NEG
 qu’ ils les remissent à **flot** (1862 *Histoire*)
 that 3P-NOM 3P-ACC put-back at stream
 “The Parisian not wishing that they put them back afloat”

Despite the possible shortcomings of our strategy to circumvent the lack of distinction between indirect object and adjuncts NP by the parsing tools, the numbers we assemble seem to confirm that bare nouns are more frequent in that function than direct objects, but less, overall, than obliques. Once again, some disparities can be noted in respect of text genre, with *Chroniques* exhibiting far fewer occurrences than the trial or procedure materials. The 15th century seems to be a crucial period, where occurrences become scarce. The residual examples come, for the most part, from set phrases (as in 36, “remettre à flot”, lit. “put back on stream”), the likes of which we still find in contemporary French.

Finally, with those numbers, we can better estimate the rate of adjuncts in the corpus, by subtracting the indirect objects from the sum of obliques (Tables 16–18, Fig. 5):

Table 16. Numbers and rates of adjunct bare nouns in *MICLE* trial material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1207 Assises Normandie</i>	21	26	44,7%
<i>1292 ANYB</i>	505	1122	31%
<i>1340 ANYB</i>	951	1623	37%
<i>1431 Proces Jeanne d'Arc</i>	304	1207	20,1%
<i>1450 Rehabilitation Jeanne d'Arc</i>	97	331	22,7%
<i>1502 Roche-Guyon</i>	360	1482	19,5 %
<i>1563 Guernesey</i>	227	326	41 %
<i>1591 Fille possedee</i>	304	1156	20,8 %
<i>1643 Proces Bavent</i>	41	156	20,8 %
Totals	2810	7429	27,4%

Table 17. Numbers and rates of adjunct bare nouns in *MICLE* procedure material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1314 Attiremens et Jugies</i>	233	432	35%
<i>1425 Style et Usage Echiquier</i>	1235	2207	39,9%
<i>1578 Terrien</i>	584	1755	24,9%
Totals	2052	4394	31,8%

Table 18. Numbers and rates of adjunct bare nouns in *Chroniques* material

	Noun without a determiner (numbers)	Noun with a determiner (numbers)	Proportion of nouns without a determiner to nouns without and with a determiner (%)
<i>1190 Kings of England</i>	74	134	35,6%
<i>1487 Cronique</i>	496	1644	23,2%
<i>1862 Histoire de la Normandie ancienne</i>	401	1992	16,8%
Totals	971	3770	20,5%

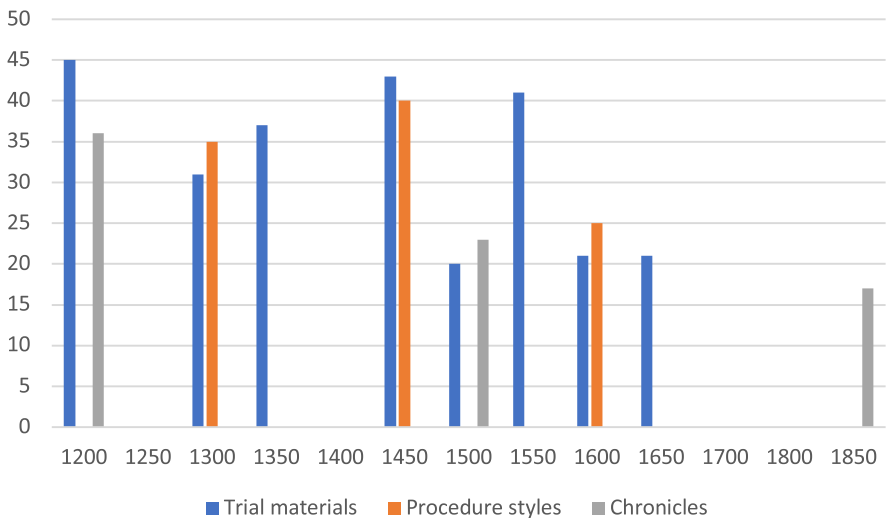


Figure 5. Evolution of adjuncts bare nouns in three sub-corpora.

From the figure, we can see that oblique bare nouns are a grammatical option all through the history of French. The 15th century, however and once again, seems to be a turning point: rates stabilize around the 20% mark, and never go lower.

5. THE COURSE OF BARE NOUNS IN MEDIEVAL FRENCH

This research confirms that bare nouns loss is impacted by syntactic function. First, in subject position, bare nouns are used at around 5% in the earliest period, to be subsequently found at around 1%, and disappear at the beginning of the 16th century. Object bare nouns go from 10% to 3%, with variation between 12% and 5% in the period between 1430 and 1550. Finally, obliques decline from a rate of nearly 40% until 1430 to around the 20% that subsist in the final period. A closer inspection of obliques shows that adjuncts go from 45% to 20% of bare nouns, and indirect objects from 40% to nil, the latter presenting a more pronounced decline: variations between texts is expected to reflect the involvement of formulaic sequences.²⁷

There are some unexpectedly high rates for some texts, but little variance between text types, an indication that comparison of results from different sub-corpora can be useful in tracking change (Amatuzzi *et al.* 2020). Again, the numbers provided by non-literary texts are lower than what is reported for literary material, because the former are presumably closer to the ongoing change, as they are not concerned with the stylistic and high register considerations found in the latter.

Syntactic function thus defines the order in which bare noun decline is occurring in the history of French as documented by these corpora. The order from subject to object to oblique (and to direct object to indirect object to adjuncts) is reminiscent of the generalization proposed by Keenan and Comrie (1977). In their work on the crosslinguistic organization of relative clause markers, they put forward an Accessibility Hierarchy.

(37) Accessibility Hierarchy:

SU > DO > IO > OBL > GEN > OCOMP

This implicational map supposes that the functions to the left are more accessible than the function to the right of the scale. The more accessible functions are more likely to be represented by a specific relative marker in any language. There are more languages with a subject marker (such as French *qui*) than with a genitive one (such as one use of French *dont*), and more with genitive than with comparative (which is ungrammatical in French). A language with relative markers in a less accessible function will comprise markers with the more accessible functions. Accessibility has also been shown to predict frequency of markers in individual languages by acquisitional studies (see the overview by Lau and Tanaka 2021). Larrivé (2020)

²⁷A reviewer points out that indirect objects bare nouns disappear more quickly than direct objects bare nouns, which is unexpected given the Hierarchy. They suggest that this may be due to indirect objects involving dative complements associated with [+human] nouns, which may tend to have a more referential interpretation than other nouns. This is an interesting avenue for future work.

documents that the frequency of French clitics is correlated to their position on the Hierarchy, and that subject clitics are more frequent than objects in spoken and written contemporary French, and objects more frequent than obliques. Regarding determiners, the Accessibility Hierarchy also helps make sense of their historical progression, and the correlative loss of bare nouns, since determiners are the most frequent in the most accessible positions at any one time, and the more accessible position see the earlier decline of bare nouns as compared to the less accessible ones.

As it set out to do, this article has provided extensive documentation on one factor correlative to the emergence of determiners. We now speculate on the question that arises that is why that should be. Why should the rate of nouns with a determiner be greater with more accessible functions? We could speculate that this relates to referentiality, that is the property of an expression of denoting an entity that exists in the common ground. Let us assume that in a language where both options are available, a noun with a determiner is more likely to be referential than the same noun without one, and that more referential nouns are found in more accessible positions, then we expect the more accessible positions to host more (common) nouns with determiners and fewer bare nouns.

The supposition that subjects host more referential nouns is supported by an early piece of work on English by Bock and Warren (1985). They showed that more referential nouns (which they call “imageable”) tend to be realized in higher functions on the Accessibility Hierarchy, and that the “more imageable” nouns occur most in subject position. Recent confirmation has been provided for one component of “imagibility” that is animacy by Thuilier *et al.* (2021) and Chen *et al.* (2022). This connection between referentiality of a noun and syntactic accessibility is of course exactly why more (definite) determiners would be found in more accessible function.²⁸

How could the supposed connection between referentiality and functions be supported from a diachronic point of view? A clue is provided by Simonenko and Carlier (2020a); they note that subjects are “at least 2–3 times more likely than objects to occur with a definite, possessive or demonstrative determiner” (2020a: 222). Assuming that definite, possessive and demonstrative determiners mark referentiality, and that NPs are more referential in more accessible function, we should then find a progression of these determiners in the more accessible functions through time. In other words, there should be more definite determiners in subject position than in object, and more in object than in oblique, at any period, and at all periods. This is what we sought to explore, using the *Chroniques* corpus. Extraction of indefinite determiners on the one hand and of definites, demonstratives and possessives on the other allowed the ratio of the latter to be calculated. We would expect the ratio of indefinites to definites to be lower with subjects than with objects, and with object than with obliques. The numbers for each function are provided below (Tables 19–21):

²⁸Whether determiners historically first emerge in more accessible functions would be an interesting issue, that cannot be pursued for French as determiner emergence precedes the first attested texts.

Table 19. Ratio of indefinite to definite determiners in subject position

	Noun with an indefinite determiner (numbers)	Noun with a definite, possessive and demonstrative determiner (numbers)	Proportion of indefinite determiners relative to definite (ratio)
<i>1190 Kings of England</i>	4	74	5%
<i>1487 Cronicque</i>	49	896	5%
<i>1862 Histoire de la Normandie ancienne</i>	86	1014	8%
Totals	139	1984	7%

Table 20. Ratio of indefinite to definite determiners in object position

	Noun with an indefinite determiner (numbers)	Noun with a definite, possessive and demonstrative determiner (numbers)	Proportion of indefinite determiners relative to definite (ratio)
<i>1190 Kings of England</i>	22	105	17%
<i>1487 Cronicque</i>	102	655	13%
<i>1862 Histoire de la Normandie ancienne</i>	237	1212	16%
Totals	361	1972	15%

Table 21. Ratio of indefinite to definite determiners in oblique position

	Noun with an indefinite determiner (numbers)	Noun with a definite, possessive and demonstrative determiner (numbers)	Proportion of indefinite determiners relative to definite (ratio)
<i>1190 Kings of England</i>	27	107	20%
<i>1487 Cronicque</i>	226	1736	12%
<i>1862 Histoire de la Normandie ancienne</i>	318	1953	14%
Totals	571	3796	13%

Using the same strategy as before, rates were produced for indirect objects and adjuncts (Tables 22 and 23):

Table 22. Ratio of indefinite to definite determiners in indirect object position

	Noun with an indefinite determiner (numbers)	Noun with a definite, possessive and demonstrative determiner (numbers)	Proportion of indefinite determiners relative to definite (ratio)
<i>1190 Kings of England</i>	2	10	20%
<i>1487 Cronique</i>	35	262	13,4%
<i>1862 Histoire de la Normandie ancienne</i>	18	193	9,3%
Totals	55	465	11,8%

Table 23. Ratio of indefinite to definite determiners in adjunct position

	Noun with an indefinite determiner (numbers)	Noun with a definite, possessive and demonstrative determiner (numbers)	Proportion of indefinite determiners relative to definite (ratio)
<i>1190 Kings of England</i>	25	97	25,8%
<i>1487 Cronique</i>	191	1474	13%
<i>1862 Histoire de la Normandie ancienne</i>	300	1760	17%
Totals	516	3331	15,5%

The data, which is in keeping with those of the other two corpora, confirms that subjects host fewer nouns with an indefinite determiner than objects; but it does not indicate that objects have fewer indefinites than obliques. Distinction between indirect objects and adjuncts is on par with the hypothesis, as the rates for adjuncts are higher than those for indirect objects, but indirect objects do not clearly evidence higher rates than objects. Note also that there is little progression for subject and object through time, and an actual decrease through time in the ratio of indefinite determiners with obliques. Clearly, more work is required around referentiality, syntactic function and determiner expression (or lack thereof). Where there is progression is in the proportion of definites, possessives and demonstratives as compared to all noun phrases, with and without determiner. For subjects, the proportions go from 78% (74/95) to 93% (896/1058) to 98% (1014/1040); and for obliques, from 46% (107/233) to 66% (1736/2637) and 72% (1953/2726). For indirect objects, proportions go from 29% (10/35) to 46% (262/569) to 45% (193/426); and for adjuncts, proportions go from 49% (97/198) to 71% (1474/2068) to 77% (1760/2300). For objects however, the proportion is stable, at around 80%. So, there is through time a growth in definite determiner use, by function, mirroring that of bare noun loss. The progress of definite determiners in more accessible functions would thus account for the loss of bare nouns along the Accessibility Hierarchy, under the assumptions presented and partially supported here by a

preliminary investigation that more accessible functions host more referential NPs that have more definite determiners.²⁹

6. CONCLUSIONS

This article seeks to make sense of the evolution of bare nouns in the history of French. On the basis of extensive documentation, we show that the curve of decline of bare nouns is determined by function, confirming suggestions by Simonenko and Carlier (2020a) and Carlier and Lamiroy (2018a; 2018b). The loss of subject bare nouns around 1500, the decrease of object bare nouns to 3%, and the reduction of obliques to 20% follow the path drawn by the Accessibility Hierarchy. Established for the investigation of relative markers by Keenan and Comrie (1977), the Hierarchy proposes that subjects are more accessible than objects, and object more than obliques.³⁰

The question that arises is why accessible functions should encourage greater rates of use of determiners. The speculation that we have proposed is that more accessible functions are characterized by greater referentiality, and that referentiality is more likely to be marked by a definite determiner. The data examined evidence that definite determiners (inclusive of possessives³¹ and demonstratives) actually see their proportion of use increase through time, being greatest with subjects, then objects, and then obliques. Thus, independent evidence shows that subjects are more referential than objects than obliques, and thus lose more quickly bare nouns, which tend to be less referential. Correlation between definiteness and syntactic position was already explored in other languages, in Malagasy (Paul 2009) or in Dutch (Rullmann 1989) for instance, and French seems to follow the same trend throughout its history.

Our study has shown that it pays to use calibrated corpora of non-fictional prose, and to compare data emerging from them. A baseline can thus be established to get closer to the effective change in the immediate competence of speakers. Outliers can be compared to results from other texts (or text types), to assess whether the figures they provide are representative of a trend in the language community, or whether they are down to individual texts or text types: the latter case seems to apply to the high levels of object bare nouns in the 1431 and 1450 trials, and to the high level of oblique bare nouns in procedure texts.

²⁹Note that this is contradiction to the idea put forward by Foulet that loss of bare nouns is due to their replacement by indefinite determiners.

³⁰As stated earlier, the annotation software does not distinguish between indirect objects and adjuncts, both being considered as “obliques”. As such, our definition of oblique NPs is broader than the one used by Keenan and Comrie, where *oblique* designates “arguments of the main predicate, as *the chest* in *John put the money in the Chest* rather than ones having a more adverbial function like *Chicago* in *John lives in Chicago*” (1977: 66). Nonetheless, we feel that this difference in scope does not overly affect the broad tendencies documented here, as shown by the figures emerging when we separate indirect objects and adjuncts. Analyses based on corpora distinguishing those would be desirable from future studies.

³¹A reviewer notes that “roughly until the end of Old French period (long) possessive pronouns are compatible with both indefinite and definite determiners (Simonenko and Carlier 2020b), suggesting that they are not definite” Fortunately, that literary feature does not figure in our corpora.

Suggestions for future research relate to method and replication. It would be nice to be able to separate indirect objects and adjuncts, which with the used parser are bunched together as obliques, although this article has proposed a strategy to do this that may possibly be generalized. Equally desirable would be to have access to more reliably annotated, non-fictional prose through time. Also, whether the pattern identified here applies to other languages, as we would expect, is certainly worthy of examination. It is to be hoped that the patterns documented here can be tested by future studies. Finally, the great question that needs to be probed is why determiners emerge at all, and under what conditions they become compulsory in some languages, but not in others.

REFERENCES

- Abeillé, Anne and Danièle Godard.** (2021). *La Grande Grammaire du Français*. 2 volumes. Arles : Actes Sud.
- Amatuzzi, Antonella, Wendy Ayres-Bennett, Annette Gerstenberg, Lene Schøsler and Carine Skupien-Dekens.** (2020). Changement linguistique et périodisation du français (pré)classique : deux études de cas à partir des corpus du RCFC. *Journal of French Language Studies* 30(3), 301–326.
- Anscombre, Jean-Claude.** (199b). La détermination zéro: quelques propriétés. *Langages* 102, 103–124.
- Anscombre, Jean-Claude.** (1991a). L'article zéro sous préposition. *Langue Française* 91, 24–39.
- Beyssade, Claire.** (2011). Bare nouns in predicate position in French. S. Pogodalla, M. Quatrini and C. Retoré (eds), *Logic and Grammar. Lecture Notes in Computer Science*. Berlin: Springer. 1–16.
- Bock, J. Kathryn and Richard K. Warren.** (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21(1), 47–67.
- Buridant, Claude.** (2019). *Grammaire nouvelle de l'ancien français*. Paris: Sedes.
- Carlier, Anne and Beatrice Lamiroy.** (2018a). The emergence of the grammatical paradigm of nominal determiners in French and in Romance: Comparative and diachronic perspectives. *Canadian Journal of Linguistics* 63(2), 141–166.
- Carlier, Anne and Beatrice Lamiroy.** (2018b). Partitive articles in Romance languages. *Oxford Research Encyclopedia of Linguistics*. <http://oxfordre.com/linguistics>
- Chaurand, Jacques.** (1991). Verbes supports et emploi ou absence de l'article. *Langages* 102, 7–22.
- Chen, Xuemei, Robert J. Hartsuiker, Merel Muylle, Mieke Sarah Slim, Chi Zhang.** (2022). The effect of animacy on structural Priming: A replication of Bock, Loebell and Morey (1992). *Journal of Memory and Language* 127. <https://doi.org/10.1016/j.jml.2022.104354>
- Combettes, Bernard.** (2001). L'émergence d'une catégorie morphosyntaxique: les déterminants du nom en français. *Linx* 45, 117–126.
- Déchainé, Rose-Marie, Monique Dufresne and Mireille Tremblay.** (2018). The trajectory of phi-features on Old French D and n. *Canadian Journal of Linguistics* 63, 167–193.
- de Swart, Henriëtte, Yoand Winter and Joostt Zwarts.** (2007). Bare nominals and reference to capacities. *Natural Language and Linguistic Theory* 25, 195–222.
- de Swart, Hentiëtte and Josst Zwarts.** (2009). Less form - more meaning: Why bare singular nouns are special. *Lingua* 119, 280–295.
- Foubat, Lucien.** (1929). *Petite syntaxe de l'ancien français*. Paris: Champion.
- Gabay, Simon, Jean-Baptiste Camps and Thibault Clérice.** (2020a). Manuel d'annotation linguistique pour le français moderne (XVI^e–XVIII^e siècles). <hal-02571190>
- Gabay, Simon, Jean-Baptiste Tanguy, Thibault Clérice, Jean-Baptiste Camps and Matthias Gille-Levenson.** (2020b). Standardizing linguistic data: method and tools for annotating (pre-orthographic) French. Proceedings of the 2nd International Digital Tools & Uses Congress (DTUC '20), Oct 2020, Hammamet, Tunisia. <https://hal.science/hal-03018381>
- Goux, Mathieu.** (2022a). L'évolution des cadres de discours dans la Coutume de Normandie. In *Le temps long: l'évolution du français dans un corpus textuel calibré. Le témoignage de la Coutume de Normandie (1250–1771)*. *Studia Linguistica Romanica*, 8, 116–130.

- Goux, Mathieu.** (ed.) (2022b). Le temps long: l'évolution du français dans un corpus textuel calibré. Le témoignage de la Coutume de Normandie (1250–1771). *Studia Linguistica Romanica*, 8.
- Goux, Mathieu and Pierre Larrivé.** (2020). Expression et position du sujet en ancien français: le rôle de la personne pronominale. *CMLF*, 78. <https://doi.org/10.1051/shsconf/20207803002>
- Goyens, Michèle.** (1994). *Émergence et évolution du syntagme nominal français*. Bern: Peter Lang.
- Grobol, Loïc and Crabbé, Benoît** (2021). Analyse en dépendances du français avec des plongements contextualisés, *Actes de la 28^e Conférence sur le Traitement Automatique des Langues Naturelles TALN-RÉCITAL 2021*. https://hal.science/hal-03223424/file/HOPS_final.pdf
- Grobol, Loïc, Prévost, Sophie and Crabbé, Benoît** (2022). Is Old French tougher to parse?, *20th International Workshop on Treebanks and Linguistic Theories, Mar 2022, Sofia, Bulgaria*. <https://hal.science/hal-03506500>
- Gross, Gaston and André Valli** (1991). Déterminant zéro et verbes supports en moyen français et en français moderne, *Langages* 102, 36–51.
- Haussalo, Teija.** (2014). *L'article zéro au XVI^e siècle : les variantes manuscrites de l'Heptaméron. Une phase de l'évolution de la langue française*. Helsinki: Publications romanes de l'Université de Helsinki.
- Heiden, Serge, Magué, Jean-Philippe and Pincemin, Bénédicte** (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement, *JADT 2010 : 10th International Conference on the Statistical Analysis of Textual Data*, http://halshs.archives-ouvertes.fr/docs/00/54/97/79/PDF/Heiden_al_jadt2010.pdf
- Keenan, Edward and Bernard Comrie.** (1977). Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* 8(1), 63–99.
- Landheer, Ronald** (1994). Présentation. *Langue française* 101, 3–11.
- Larrivé, Pierre.** (2020). Le registre comme continuum. L'interaction entre proximité de l'échange et hiérarchie d'accessibilité. Franck Neveu, Bernard Harmegnies, Linda Hriba, Sophie Prévost and Agnes Steuckardt (dirs), *Congrès Mondial de Linguistique Française 2018*. SHS Web of Conferences.
- Lau, Elaine and Nozomi Tanaka.** (2021). The subject advantage in relative clauses: A review. *Glossa: a journal of general linguistics* 6(1): 34.
- Lauwers, Peter.** (2012). Does French really have bare singular subject NPs? On semantic reconstruction partly based on (cata)phoric ellipsis. *Neuphilologische Mitteilungen* 113(1), 3–40.
- Lauwers, Peter.** (2014). Copular constructions and adjectival uses of bare nouns in French: a case of syntactic recategorization? *Word* 58(2), 89–122.
- Lauwers, Peter D.** 2011. The modification of predicative bare nouns in French: a functional analysis. *Transactions of the Philological Society* 109, 12–40.
- Lorian, Alexandre.** (1973). *Tendances stylistiques dans la prose narrative française du XVI^e siècle*. Paris: Klincksieck.
- Marchello-Nizia, Christiane.** (1979). *Histoire de la langue française au XIV^e et XV^e siècles*, Paris: Bordas.
- Marchello-Nizia, Christiane.** (1999). *Le français en diachronie: douze siècles d'évolution*. Gap: Ophrys.
- Marchello-Nizia, Christiane, Bernard Combettes, Sophie Prévost and Tobias Scheer** (eds). (2020). *Grande grammaire historique du français*. Berlin: De Gruyter.
- Martineau, France.** (2008). Un corpus pour l'analyse de la variation et du changement linguistique. *Corpus 7*. <https://doi.org/10.4000/corpus.1508>
- Märzhäuser, Christina.** (2013). Coordinated bare nouns in French, Spanish and European Portuguese. Johannes Kabatek and Albert Wall (eds), *New Perspectives on Bare Noun Phrases in Romance and beyond*. Amsterdam: Benjamins. 283–300.
- Paul, Ileana.** (2009). On the Presence versus Absence of Determiners in Malagasy. Jila Ghomeshi, Ileana Paul, and Martina Wiltschok (eds), *Determiners: Universals and Variation*. Amsterdam: Benjamins. 215–242. <https://ir.lib.uwo.ca/frenchpub/65/>
- Pinkster, Harm.** (2015). *The Oxford Latin Syntax. Volume 1: The Simple Clause*. Oxford: Oxford University Press.
- Prévost, Sophie** (2015). Diachronie du français et linguistique de corpus: une approche quantitative renouvelée. *Langages*, 197(1), 23–45.
- Riegel, Martin, Jean-Christophe Pellat, and René Rioul** (2014). *Grammaire méthodique du français*. Paris: Presses Universitaires de France.
- Roodenburg, Jasper.** (2004). French bare arguments are not extinct: The case of coordinated bare nouns. *Linguistic Inquiry* 35(2), 301–313.

- Rullmann, Hotze.** (1989). Indefinite Subjects in Dutch. *University of Massachusetts Occasional Papers in Linguistics* 12, article 14. <https://scholarworks.umass.edu/umop/vol12/iss1/14>
- Saikali-Sleiman, Dana.** (2006). *L'absence d'article en français : étude diachronique et sémantique*. PhD Thesis, Paris 10, Annie Bertin (dir.).
- Sakari, Ellen.** (1988). La détermination du nom en moyen français : l'article défini et ses « substituts » opposés au marquant zéro dans les *Mémoires* de Philippe de Commynes. U. Jokinen and P. Sihvonen-Hauteceur (eds), *Approches du moyen français*. Jyväskylä: University of Jyväskylä, 141–178.
- Simonenko, Alexandra and Anne Carlier.** (2020a). Givenness marking in a mixed system: Constituent order vs. Determiners. Remus Gergel and Jonathan Watkins (eds.), *Quantification and scales in change*. Berlin: Language Science Press, 199–228.
- Simonenko, Alexandra and Anne Carlier.** (2020b). Between demonstrative and definite: A grammar competition model of the evolution of French I-determiners. *Canadian Journal of Linguistics* 65(3), 393–437.
- Skrzypek, Dominika, Alicja Piotrowska and Rafał Jaworski.** (2021). *The diachrony of definiteness in North Germanic*. Leiden and Boston: Brill.
- Sornicola, Rosanna.** (2009). Continuités textuelles et discontinuités catégorielles : le cas de l'article roman. *Travaux de linguistique* 59(2), 25–46.
- Thuilier, Juliette, Margaret Grant, Benoît Crabbé, and Anne Abeillé,** (2021) Word order in French: the role of animacy, *Glossa: a journal of general linguistics* 6(1): 55. doi: <https://doi.org/10.5334/gjgl.1155>
- Vigier, Denis.** (2015). *En, dans, dedans* prépositions au XVI^e siècle. Approche statistique et combinatoire. *Le français moderne* 83(2), 230–247.
- Vigier, Denis.** (2017). L'évolution des usages des prépositions *en, dans, dedans* entre le XVI^e et le XX^e siècle : approche distributionnelle sur corpus outillé. *Discours*, 21, <https://doi.org/10.4000/discours.937>

Cite this article: Larrivée P and Goux M (2024). The evolution of bare nouns in the history of French. The view from calibrated corpora. *Journal of French Language Studies*, 1–28. <https://doi.org/10.1017/S0959269524000061>